



Contribution ID: 52

Type: **not specified**

Persistent Homology for Market Basket Analysis

Monday, 17 May 2021 14:45 (20 minutes)

In the last years, the possibility of improving the analysis of data by capturing intrinsic relations has proven to be a flourishing research direction in data analysis. Graphs and higher-order structures are more and more often associated to data in order to infer qualitative knowledge, possibly independently from the data embedding representation. In this direction, topological data analysis is one of the emerging research fields. Its aim is that of studying data under the lens of topology, a branch of mathematics dealing with shape properties that are invariant under continuous deformations. In the case of customer behaviour analysis, much of the potentiality of these new research trends is still to be investigated.

In this work, we want to present a preliminary investigation of persistent homology, a standard tool in topological data analysis, applied to market basket analysis. To do this, we will present the construction of a filtered simplicial complex to represent the purchased item sets in a relational consistent way. Then, we will highlight correspondences between the presence of topological features along the filtered simplicial complex, such as loops and cavities, and standard metrics in market basket analysis, such as confidence and lift measures. Some preliminary comparisons on real datasets will be presented.

The authors acknowledge the connection of this work with ELBA, an ERASMUS+ project aiming at the establishment of training and research centers on Big Data Analysis in Central Asia (<https://elba.famnit.upr.si>).

Primary authors: SCARAMUCCIA, Sara (Politecnico di Torino); Prof. FONTANA, Roberto (Politecnico di Torino)

Presenter: SCARAMUCCIA, Sara (Politecnico di Torino)

Session Classification: Statistical models and applications

Track Classification: Data Science in Process Industries